

# A Genomic Score Prognostic of Outcome in Trauma Patients

H Shaw Warren,<sup>1\*</sup> Constance M Elson,<sup>2\*</sup> Douglas L Hayden,<sup>2\*</sup> David A Schoenfeld,<sup>2</sup> J Perren Cobb,<sup>3</sup> Ronald V Maier,<sup>4</sup> Lyle L Moldawer,<sup>5</sup> Ernest E Moore,<sup>6</sup> Brian G Harbrecht,<sup>7</sup> Kimberly Pelak,<sup>2</sup> Joseph Cuschieri,<sup>4</sup> David N Herndon,<sup>8</sup> Marc G Jeschke,<sup>8</sup> Celeste C Finnerty,<sup>8</sup> Bernard H Brownstein,<sup>3</sup> Laura Hennessy,<sup>4</sup> Philip H Mason,<sup>9</sup> Ronald G Tompkins,<sup>9</sup> and The Inflammation and the Host Response to Injury Large Scale Collaborative Research Program

<sup>1</sup>Infectious Disease Unit, Massachusetts General Hospital, Charlestown, Massachusetts, United States of America; <sup>2</sup>Biostatistics Center, Massachusetts General Hospital, Boston, Massachusetts, United States of America; <sup>3</sup>Department of Surgery, Washington University, St. Louis, Missouri, United States of America; <sup>4</sup>Department of Surgery, Harborview Medical Center, University of Washington, Seattle, Washington, United States of America; <sup>5</sup>Department of Surgery, University of Florida College of Medicine, Gainesville, Florida, United States of America; <sup>6</sup>Department of Surgery, University of Colorado, Denver Health Medical Center, Denver, Colorado, United States of America; <sup>7</sup>Department of Surgery, University of Louisville, Louisville, Kentucky, United States of America; <sup>8</sup>Department of Surgery, University of Texas Medical Branch, Shriners Burn Hospital, Galveston, Texas, United States of America; <sup>9</sup>Department of Surgery, Massachusetts General Hospital, Boston, Massachusetts, United States of America

Traumatic injuries frequently lead to infection, organ failure, and death. Health care providers rely on several injury scoring systems to quantify the extent of injury and to help predict clinical outcome. Physiological, anatomical, and clinical laboratory analytic scoring systems (Acute Physiology and Chronic Health Evaluation (APACHE), Injury Severity Score (ISS)) are utilized, with limited success, to predict outcome following injury. The recent development of techniques for measuring the expression level of all of a person's genes simultaneously may make it possible to develop an injury scoring system based on the degree of gene activation. We hypothesized that a peripheral blood leukocyte gene expression score could predict outcome, including multiple organ failure, following severe blunt trauma. To test such a scoring system, we measured gene expression of peripheral blood leukocytes from patients within 12 h of traumatic injury. cRNA derived from whole blood leukocytes obtained within 12 h of injury provided gene expression data for the entire genome that were used to create a composite gene expression score for each patient. Total blood leukocytes were chosen because they are active during inflammation, which is reflective of poor outcome. The gene expression score combines the activation levels of all the genes into a single number which compares the patient's gene expression to the average gene expression in uninjured volunteers. Expression profiles from healthy volunteers were averaged to create a reference gene expression profile which was used to compute a difference from reference (DFR) score for each patient. This score described the overall genomic response of patients within the first 12 h following severe blunt trauma. Regression models were used to compare the association of the DFR, APACHE, and ISS scores with outcome. We hypothesized that patients with a total gene response more different from uninjured volunteers would tend to have poorer outcome than those more similar. Our data show that for measures of poor outcome, such as infections, organ failures, and length of hospital stay, this is correct. DFR scores were associated significantly with adverse outcome, including multiple organ failure, duration of ventilation, length of hospital stay, and infection rate. The association remained significant after adjustment for injury severity as measured by APACHE or ISS. A single score representing changes in gene expression in peripheral blood leukocytes within hours of severe blunt injury is associated with adverse clinical outcomes that develop later in the hospital course. Assessment of genome-wide gene expression provides useful clinical information that is different from that provided by currently utilized anatomic or physiologic scores.

© 2009 The Feinstein Institute for Medical Research, [www.feinsteininstitute.org](http://www.feinsteininstitute.org)

Online address: <http://www.molmed.org>

doi: 10.2119/molmed.2009.00027

## INTRODUCTION

Trauma is now the number one health care cost in the United States (1) and a

major health priority throughout the world (2,3,4). Traumatic injury frequently leads to subsequent development of in-

fections, sepsis, and multiple organ dysfunction (5,6), resulting in high morbidity and mortality. Multiple organ dysfunction syndrome (MODS) is associated with extended intensive care unit (ICU) and hospital stays and poor prognosis; its pathophysiology remains an area of intense study (7). The broad range of patient outcomes following similar traumatic insults has spurred the development of trauma-specific scoring systems in an attempt to stratify and predict clinical

---

\*HSW, CME, and DLH are coequal, first authors.

**Address correspondence and reprint requests to** Douglas Hayden, MGH Biostatistics Center, 50 Staniford Street #560, Boston, MA 02114. Phone: 617-726-3720; Fax: 617-724-9878; E-mail: [dhayden@partners.org](mailto:dhayden@partners.org); H. Shaw Warren, Infectious Disease Unit, Massachusetts General Hospital East, 149 13th Street, Charlestown, MA 02129. Phone: 617-726-5774; Fax: 617-726-5411; E-mail: [swarren1@partners.org](mailto:swarren1@partners.org).

Submitted March 30, 2009; Accepted for publication April 7, 2009; Epub ([www.molmed.org](http://www.molmed.org)) ahead of print April 10, 2009.

cal trajectories. Commonly used scoring systems are a composite of organ-based physiologic, hematologic, and chemical measurements or are anatomic-specific injury scores. These systems include the Acute Physiology and Chronic Health Evaluation (APACHE) II or APACHE III scores (8,9) and the Injury Severity Score (ISS) (10,11,12).

The most common outcomes that are used for the study of severe blunt trauma include 28-day all-cause mortality and measures of morbidity. With increasing survival rates following blunt trauma, outcomes have focused more specifically on evidence of specific organ dysfunction (such as duration of ventilator dependence) as well as infectious complications and lengths of stay in the ICU and in the hospital. To evaluate overall clinical state, composite scores of multiple organ dysfunction such as the Marshall and Denver organ failure scores (13,14,15,16) often are utilized.

The recent development of high-throughput microarray systems that can survey the mRNA abundance across the entire human genome has paved the way for studies exploring the association between the human transcriptome and human health and disease. We hypothesized that early changes in gene expression from peripheral blood leukocytes, after severe traumatic injury, might be used to predict outcomes in individual subjects. Our overarching hypotheses were that the early genomic response reflects activation of innate immunity and inflammation that determines the ultimate clinical outcome of the patient, and that variations in gene expression would be associated with different clinical trajectories. To test these hypotheses, genome-wide gene expression data derived from blood leukocyte samples from trauma patients in six United States medical centers were used to calculate a composite difference from reference (DFR) score for each patient. This score was developed to summarize the changes in gene expression across the entire genome by using a single number

which represented the aggregate difference of each patient's gene expression from a healthy genomic reference profile. In 158 trauma patients, we computed the DFR score and its association with selected clinical outcomes, including multiple organ failure, ICU and hospital length of stay, ventilator dependence, infectious complications, and mortality. The same associations with outcome were reexamined after controlling for degree of injury as measured by the clinical scores. The results suggest that assessment of gene expression across the entire genome provides prognostic information for patients with severe trauma as well as, and independently from, presently available clinical scoring systems.

## MATERIALS AND METHODS

### Subject Recruitment and Data Collection

The 158 patients used for this study were recruited between November 2003 and January 2005 at six participating trauma centers (17). All patients had suffered severe blunt trauma but did not have severe brain injury defined as Glasgow Coma Score less than nine with abnormal CT scan of the head. Patients were expected to survive beyond 24 h, ranged from 16 to 55 years in age, and were 64% male. Patients were analyzed in total and in two subgroups: the first 79 and subsequent 79 patients enrolled in the study.

The 26 control subjects were recruited at Harborview Medical Center and the University of Texas Medical Branch-Galveston. Inclusion criteria consisted only of being in apparent good health. The controls ranged from 18 to 41 years old and 53% were male. Controls included four young adults ages 18 to 20 who had suffered serious burn injuries 3 to 6 years previously; results were similar when they were excluded from the analysis.

The Institutional Review Board of each participating center approved the study, and written informed consent was ob-

tained from all patients or their legal next of kin and from all volunteer control subjects.

For each traumatized patient, peripheral blood samples were taken within 12 h of injury. Total blood leukocytes from patients and controls were processed according to protocols published previously (18,19). Very briefly, EDTA-anticoagulated whole blood was collected and processed at room temperature within 1 h of sample collection. Blood was centrifuged at 400g for 8 min and the resulting plasma removed. The residual red and white blood cell fraction was subsequently diluted twenty-five-fold with a lysis buffer (EL Buffer, Qiagen, Valencia, California, USA) and placed on ice for 15 min. Thereafter, the sample was centrifuged at 400g at 4° C for 8 min and the supernatant removed. The leukocyte pellet was washed a second time with 15 mL of room temperature EL Buffer. The sample was centrifuged a second time, and the supernatant removed. The cell pellet was dried and lysed with RLT buffer (Qiagen), and total cRNA extracted according to the manufacturer's instructions (Qiagen RNeasy Midi Kit Cat 75142). The resulting cRNA was hybridized onto Affymetrix GeneChips to generate the gene expression data used in this study.

Clinical data were obtained by trained study nurses, entered into a study-wide database and checked for consistency and accuracy by a data manager. Modified Marshall and Denver scores were used to assess MODS in the ICU. The modified Marshall score used in this study omitted the neurologic component and was the sum of the five remaining component scores. The Denver score was the sum of four component scores, including a cardiac component determined by the level of inotropes administered. (20) The APACHE II score (8), the Injury Severity Score (ISS) (10), and the anatomical injury score (AIS) (10) were calculated as described. Definitions of infectious and noninfectious complications were as described (20).

Clinical outcomes that occurred in the ICU and within 28-d posttrauma were

recorded. Primary outcomes of interest included 28-d all-cause mortality, presence of MODS, nosocomial infections or complications, and overall length of ICU and hospital stay. The presence of MODS was arbitrarily described as either attaining a modified Marshall score of six or higher, or attaining a Denver score of four or higher during the 28 d post trauma period.

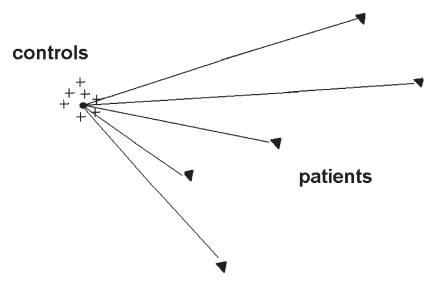
### Data Analysis and Calculation of DFR Score

Genome-wide gene expression analysis was obtained from peripheral blood leukocyte samples, using cRNA hybridized to an Affymetrix HU133 Plus 2.0 GeneChip (according to the manufacturer's recommendations.) On this chip, there are a total of 54,613 probe sets whose scanned values were normalized and modeled using dChip (21) software, as we have described (20). The model consists of a fixed set of parameters to create probe set expression estimates as weighted sums of individual probe levels. The resulting 54,613 probe set expressions constituted a gene expression profile for each subject.

Gene expression profiles were computed for the healthy control subjects. A healthy reference profile was constructed from these control profiles as follows: for every probe set ( $i$ ), the mean ( $M_i$ ), and variance ( $V_i$ ) of the probe set expressions within the control group were computed. The collection of the  $M_i$  over all the probe sets formed the healthy reference profile. The healthy reference profile might be conceptualized graphically as the center of a group of gene expression profiles from healthy control subjects (Figure 1).

From a healthy reference profile, a DFR score for each patient was created by squaring the difference between the expression in the patient and the expression in the reference profile for each probe set, scaling this by the control group variance, and summing over all 54,613 probe sets as described by the formula:

$$\text{DFR} = \ln \sum_{\text{probe sets}} \frac{(e_i - M_i)^2}{V_i}$$



**Figure 1.** Conceptual representation of genomic DFR score. The DFR score can be visualized as the distance from the patient genomic profile (▼) to a healthy reference profile (•).

where  $e_i$  is the patient's expression level and  $M_i$  and  $V_i$  are the control group mean and variance for the  $i^{\text{th}}$  probe set. Division by the control variance is a rescaling that prevents the DFR score from being dominated by genes that are inherently more variable or more highly expressed. The natural logarithm was applied to make the distribution of the resulting DFR more symmetric over the patient population. A patient's DFR genomic score might be visualized graphically as the distance between the patient's gene expression profile and the profile obtained from the healthy reference subjects (see Figure 1).

DFR, APACHE, and ISS scores represent genomic, physiologic, or anatomic assessments of the patient status in the first 24 h after severe traumatic injury. Pearson correlations between each of these scores and baseline and injury severity variables were computed. Univariate regression models tested the association between outcomes of interest and each of the patient scores. For each score, linear regression was used to model continuous outcomes and logistic regression was used to model event outcomes. To assess how much information about patient prognosis was contained in the genomic score beyond that already contained in the clinical scoring system, the association of DFR score with outcome was tested while controlling for the effect of APACHE II score. This analysis

was repeated using the initial ISS as the clinical score.

### Analysis of Patient Subgroups

The concept of the DFR was based on the theoretical consideration that the difference between a patient's gene expression and that of an uninjured person would measure the extent of the patient's immunological aberration. The definition was determined prior to testing any patients. This notion was first tested on 79 patients and 10 controls recruited in the first year of the study. We retested our results on the next cohort of 79 subjects and 16 controls recruited in the second year of the study. We present the results of the combined sample (158 patients, 26 controls) and both subgroups. Pooling the data was valid statistically because the DFR score had been defined and fixed independent of clinical patient data.

### Web Deposition of Data

Data in this study have been deposited in the Gene Expression Omnibus (GEO) site (<http://www.ncbi.nlm.nih.gov/geo/>), accession number GSE11375.

### RESULTS

Table 1 presents baseline, injury, and outcome clinical data for the entire 158 severely injured patients.

Pearson correlations of the DFR, APACHE II, and ISS patient scores with the baseline and injury severity variables listed in Table 1 indicated that DFR and APACHE scores were correlated significantly with volume of blood transfused during the initial 12 h, the worst base deficit during the initial 12 h, and maximum Anatomic Injury score. The DFR score also was correlated with age and hypotension. The APACHE II score was correlated with ventilator status on admission. The ISS score was correlated only with the maximum Anatomic Injury score (10) over all body regions. No other correlations were significant.

The relationship of patient scores to outcome was analyzed by univariate regression and is shown in Table 2. The

**Table 1.** Patient baseline characteristics, injury severity, and outcomes.

Baseline	
Age, mean (SD), y	34.0 (11.2)
Over 40 y old, N (%)	51 (32%)
Male, N (%)	101 (64%)
No comorbidities, N (%)	45 (28%)
Hypotension, medicated, N (%)	11 (7%)
Smoker, N (%)	63 (40%)
Chronic alcohol abuse, N (%)	24 (15%)
Injury severity	
Admitted on ventilator, N (%)	88 (56%)
APACHE score, mean (SD)	27.3 (6.0)
Injury Severity Score, mean (SD)	31.1 (13.5)
Maximum anatomic injury score, mean (SD)	4.0 (0.9)
Blood transfused, 0–12 h, mean (SD), mL	2460 (2050)
Worst base deficit, 0–12 h, mean (SD)	–10.0 (4.6)
Outcome	
Days in ICU, mean (SD)	13.2 (11.5)
Days on ventilator, mean (SD)	9.7 (8.5)
Hospital length of stay, mean (SD)	24.8 (18.3)
Maximum Marshall score, mean (SD)	5.5 (2.8)
Maximum Denver score, mean (SD)	2.1 (1.9)
MODS (Marshall criterion: Marshall score $\geq 6$ ), N (%)	67 (42%)
MODS (Denver criterion: Denver score $\geq 4$ ), N (%)	24 (15%)
Complications, N (%)	82 (52%)
Nosocomial infections, N (%)	87 (55%)
Death, N (%)	7 (4%)

DFR, APACHE II, and ISS scores were associated strongly with almost all outcomes, and their degrees of association were generally comparable.

To evaluate whether a patient's DFR score provided additional information beyond that provided by the existing physiologic scores, multivariate models tested for association of outcome with DFR score after controlling for the effects of the APACHE II or ISS scores. Table 3 presents results obtained by regressing outcome on APACHE and DFR scores. These multivariate models were statistically significant for all outcomes shown, except for 28-d all-cause mortality. For any fixed level of APACHE score, the genomic DFR score showed significant positive association with the duration of ventilator dependence, ICU and hospital length stay, maximum Marshall and Denver scores, and with the presence of MODS defined by the Denver scoring system. The DFR score was associated only weakly with the development of nosocomial infections and with MODS

defined using the Marshall scoring system. Similar results were obtained when

we controlled for ISS, regressing outcome on ISS and DFR scores (data not shown).

Table 4 compares the association of the DFR score with clinical outcome for the two patient subgroups. The DFR scores for the two patient subgroups were derived from distinct reference profiles, each calculated using control subjects associated with the patient subgroup. For the initial 79 patients and 10 controls, there was a highly significant association of DFR with most outcomes, similar to the pooled results presented in Table 2. The association of DFR with outcome in the second set of 79 patients and 16 controls was qualitatively similar to Table 2, but generally weaker. Significant association was attained for only three outcomes and was attained weakly for an additional outcome. At present there is no easy explanation for this difference in performance in the two subgroups, although it is probably due to natural biological variability. The distribution of DFR scores in the first and second patient subgroups was similar but not identical. Both distributions were roughly normal with a mean and standard deviation

**Table 2.** Comparison of association of outcome with patient scores.

	DFR score	APACHE II score	Injury Severity Score
Outcome <sup>a</sup>			
Days in ICU	2.9 $\pm$ 0.9 (0.001)	2.9 $\pm$ 0.9 (0.001)	3.3 $\pm$ 0.9 (< 0.001)
Days on ventilator	3.0 $\pm$ 0.6 (< 0.001)	2.9 $\pm$ 0.7 (< 0.001)	2.1 $\pm$ 0.7 (0.002)
Hospital length of stay	4.4 $\pm$ 1.4 (0.002)	3.8 $\pm$ 1.4 (0.009)	3.5 $\pm$ 1.4 (0.02)
Maximum Marshall score	1.1 $\pm$ 0.2 (< 0.001)	1.1 $\pm$ 0.2 (< 0.001)	0.8 $\pm$ 0.2 (< 0.001)
Maximum Denver score	0.7 $\pm$ 0.2 (< 0.001)	0.6 $\pm$ 0.2 (< 0.001)	0.4 $\pm$ 0.2 (0.005)
Event <sup>b</sup>			
Mortality	1.9 (0.9, 3.6) (0.08)	2.4 (1.0, 5.8) (0.06)	1.9 (0.9, 4.1) (0.08)
MODS (Marshall criterion)	1.6 (1.1, 2.2) (0.008)	1.8 (1.3, 2.6) (0.001)	1.6 (1.1, 2.2) (0.006)
MODS (Denver criterion)	2.1 (1.4, 3.3) (< 0.001)	1.8 (1.1, 3.0) (0.02)	1.5 (1.0, 2.4) (0.05)
Nosocomial Infections	1.6 (1.2, 2.3) (0.006)	2.0 (1.4, 2.9) (< 0.001)	1.7 (1.2, 2.4) (0.002)
Surgical Site Infections	1.4 (1.0, 2.0) (0.08)	1.8 (1.2, 2.8) (0.009)	1.3 (0.9, 1.8) (0.24)

<sup>a</sup>The model is Outcome = Intercept +  $\alpha$  Score +  $\epsilon$ . Each table entry shows the estimated regression coefficient  $\alpha$ , its standard error, and its significance. Standardized scores were used in these models. The means and standard deviations used to standardize patient scores were : DFR 13.39 (0.68), APACHE 27.25 (6.01), ISS 31.13 (13.46).

<sup>b</sup>For event outcomes, the model is Log Odds (Event) = Intercept +  $\alpha$  Score. Each table entry shows the estimated odds ratio, its 95% CI, and its significance. The odds ratio is exp ( $\alpha$ ); it indicates the increased (or decreased) odds of the event for one unit increase in the standardized score.



**Table 3.** Multivariate regression models for association of outcome with DFR score after controlling for APACHE II score.

	APACHE II score	DFR score	Model significance
Outcome <sup>a</sup>			
Days in ICU	2.2 ± 0.9 (0.02)	2.2 ± 0.9 (0.02)	< 0.001
Days on ventilator	2.1 ± 0.7 (0.002)	2.3 ± 0.7 (< 0.001)	< 0.001
Hospital length of stay	2.6 ± 1.5 (0.08)	3.6 ± 1.5 (0.02)	0.002
Maximum Marshall score	0.7 ± 0.2 (< 0.001)	0.8 ± 0.2 (< 0.001)	< 0.001
Maximum Denver score	0.5 ± 0.2 (0.002)	0.5 ± 0.2 (< 0.001)	< 0.001
Event <sup>b</sup>			
Mortality	1.9 (0.7, 7.0) (0.19)	1.5 (0.7, 3.1) (0.32)	0.09
MODS (Marshall criterion)	1.7 (1.1, 2.4) (0.009)	1.4 (0.9, 1.9) (0.09)	< 0.001
MODS (Denver criterion)	1.4 (0.8, 2.3) (0.24)	1.9 (1.2, 3.0) (0.01)	0.001
Nosocomial infections	1.9 (1.3, 2.7) (0.001)	1.4 (0.95, 2.0) (0.09)	< 0.001
Surgical site infections	1.7 (1.0, 2.6) (0.03)	1.2 (0.8, 1.8) (0.42)	0.02

<sup>a</sup>The model is Outcome = Intercept +  $\alpha$  APACHE II +  $\beta$  DFR +  $\epsilon$ . Table entries show the estimated regression coefficients, their standard errors, and significance. The means and standard deviations used to standardize scores were given in Table 2.

<sup>b</sup>For event outcomes, the model is Log Odds (Event) = Intercept +  $\alpha$  APACHE II +  $\beta$  DFR. Each table entry shows the estimated odds ratio, its 95% CI and its significance. The odds ratio indicates the increased (or decreased) odds of the event for one unit increase in the standardized score when the other variable is held fixed.

tion of  $13.97 \pm 0.58$  in the first subgroup and  $14.46 \pm 0.61$  in the second. The two groups of patients were statistically equivalent for all the clinical indicators presented in Table 1 except for three: the

rate of admission on ventilator was greater for the second group (65% versus 46%), their complication rate was higher (62% versus 42%), and the average maximum Denver score was higher (2.4 ver-

**Table 4.** Comparison of outcome-DFR association for two patient subgroups.<sup>a</sup>

	Group 1	Group 2
Outcome <sup>b</sup>		
Days in ICU	4.8 ± 0.9 (< 0.001)	0.5 ± 1.5 (0.74)
Days on ventilator	3.8 ± 0.8 (< 0.001)	2.1 ± 1.0 (0.04)
Hospital length of stay	7.1 ± 1.9 (< 0.001)	1.0 ± 2.1 (0.65)
Maximum Marshall score	1.3 ± 0.3 (< 0.001)	0.6 ± 0.3 (0.04)
Maximum Denver score	0.7 ± 0.2 (< 0.001)	0.5 ± 0.2 (0.05)
Event <sup>c</sup>		
Mortality	2.7 (1.0, 7.1) (0.05)	1.4 (0.5, 3.9) (0.57)
MODS (Marshall criterion)	3.0 (1.6, 5.7) (< 0.001)	1.1 (0.7, 1.7) (0.78)
MODS (Denver criterion)	2.7 (1.3, 6.0) (0.006)	1.7 (0.9, 3.2) (0.08)
Nosocomial infections	1.9 (1.1, 3.2) (0.02)	1.1 (0.7, 1.7) (0.72)
Surgical site infections	1.2 (0.70, 2.4) (0.43)	1.5 (0.9, 2.5) (0.15)

<sup>a</sup>Group 1 is 79 patients with DFR score based on 10 controls. Group 2 is 79 patients with DFR score based on 16 controls. Standardized scores were used. The mean and standard deviation used to standardize patient scores were: Group 1 DFR 13.97 (0.58), Group 2 DFR 14.46 (0.61).

<sup>b</sup>The model is Outcome = Intercept +  $\alpha$  DFR +  $\epsilon$ . Each table entry shows the estimated regression coefficient  $\alpha$ , its standard error, and its significance.

<sup>c</sup>For event outcomes, the model is Log Odds (event) = Intercept +  $\alpha$  DFR. Each table entry shows the estimated odds ratio, its 95% CI, and its significance.

sus 1.7). The association of the DFR score and clinical outcomes depended strongly on the patient subgroup but only depended weakly on the control reference subgroup (data not shown).

## DISCUSSION

For decades, physicians have sought methods to stratify critically ill patients early after injury into groups that would predict their subsequent clinical trajectory and outcome. The main finding of this study is that a single genomic score derived from the genome-wide gene expression of peripheral blood leukocytes collected within 12 hours after severe blunt injury can provide useful information regarding the subsequent outcome of the injured patients.

At the present time, prediction of outcomes is best determined by scoring systems based upon anatomic patterns of injury and/or physiologic data. The ISS first was described in 1974 and was designed to quantify the extent of injury in a patient sustaining injuries to several parts of the body. It combines defined injury scoring scales from 0–5 for up to three injured regions of the body, resulting in a single severity score shown to correlate with outcome. The APACHE system of injury classification was first described in 1985. It quantifies physiologic dysfunction using the degree of abnormality of 12 physiologic variables over the first 24 hours of a patient's admission to a hospital. Whereas the first APACHE system used 34 variables, the APACHE II eliminated those variables that were rarely available or that overlapped with other variables, which helped it correlate more accurately with outcome. The APACHE II scoring system is utilized widely to assess the severity of illness and for prognosis in different settings. Both scoring systems are used to stratify and identify different patient groups for clinical studies although the scores were not designed for and have not been validated for this purpose. The APACHE II score is not available until the second day of illness, requires operator input, and may vary up to 15% de-

pending upon physician input (22,23). ISS frequently requires modification as initial occult injuries are identified, and it often does not account for multiple injuries in a single anatomic region, which may directly impact outcome.

Identification of biologically important signals from the enormous amount of data generated by genome-wide expression analyses presents a number of challenges. One approach has been to attempt to identify individual genes significantly associated with clinical outcome. However, the extraction of biological or pathophysiological meaning from very large datasets is extremely difficult when the list of differentiated genes includes thousands of entries. The approach presented here uses an analytic method that was designed *a priori* to reduce the entire patient transcriptome to a single number. There was no attempt to identify specific genes that differed between trauma patients and control subjects or between patients with different clinical outcomes. With the variance scaling, all gene information was incorporated with equal weight.

Analysis of gene expression in peripheral blood leukocytes has been used with varying degrees of success for diagnosis and outcome prediction in fields as diverse as oncology, neurodegenerative disease, autoimmune disease, surgery, and neuropsychiatric disorders. For the most part, the studies attempt to focus on identifying a limited numbers of genes containing biological information (24,25). However, outcomes such as death, healing rate, or massive inflammation are themselves summary outcomes that undoubtedly involve interplay of many genomic processes. The DFR score or similar genomic scores that reflect changes in the expression of the entire genome or selected large portions of it may better summarize complex multigene processes than analyses focused on smaller numbers of genes.

Because the DFR score was computed independently of the clinical data, model overfitting was not an issue

when analyzing association of the score with clinical outcome. This approach is different from most contemporary data mining, in which an initial training set of data is used to estimate parameters that link the properties or variables of interest, and then a second set of data is used to test the validity of the model. The concept and algorithm for the DFR score were developed without reference to the clinical state of the first 79 patients. The subsequent 79 patients showed a weaker association between the DFR score and disease than was observed in the first group. However, neither group served as a training or validation set for a model. The pooled dataset presented in Tables 1–3 provide the most precise estimates of association of the three patient scores with established clinical outcome parameters.

A major finding of the analysis was that even after controlling for the degree of physiologic dysfunction (using the APACHE II or ISS scores), higher DFR scores were associated with poor outcome. This finding strongly suggests that there is information relating to clinical outcome that is contained in the measurement of gene expression that is not captured by current scoring systems. It also raises the possibility that some combination of initial genomic and physiologic scores could provide useful prognostic information about patient outcomes that is superior to each score alone.

There are several limitations to the present study. First, the number of patients in each of the analysis groups was relatively small. The small numbers precluded sufficient power to detect association with infrequent outcomes such as death. Second, the statistical associations with some outcomes of the DFR score were weaker in the second group of 79 patients. Third, as with many studies of gene expression in peripheral blood leukocytes, the results reflect changes in gene expression in the entire population of leukocytes in peripheral blood. Different results might be obtained when exploring subpopulations of leukocytes.

Fourth, because the DFR score introduced here represents a summary of expression of the entire genome, the study does not add any information regarding the role of any individual genes in the disease processes that lead to the outcome studied. At the current time, it is not known whether there are some genes or gene groups that are primarily responsible for the significant associations with clinical state. Variants of the DFR score based on specific genes or gene families may improve on the results shown here and may be a tool to better understand pathophysiology that is specific to individual patients.

Despite these limitations, the study provides proof-of-principle that measurement of genome-wide gene expression in peripheral blood leukocytes in the first hour after massive trauma can provide useful information that cannot be obtained using current anatomic or physiological scoring systems that are obtained later in the hospital course. It does not seem surprising that early changes in the transcriptome should, in fact, predict subsequent changes in clinical status. A blood test would have some advantages over current systems because it can be obtained at admission and is independent of provider input. In addition to providing prognostic information, such a scoring system might also be helpful to identify and balance patients in clinical trials in which genomic components are not currently taken into account. Finally, because this approach provides information even when adjusted for anatomic and physiologic scores such as ISS and APACHE II, it should be possible to utilize anatomic, physiologic, and genomic data to develop a combined scoring system that performs better than any one method alone.

## ACKNOWLEDGMENTS

The magnitude of the clinical and genomic data reported here required the efforts of many individuals at participating institutions. In particular, we wish to acknowledge the supportive research envi-

ronment created and sustained by the participants in the Glue Grant Program: Henry V Baker, University of Florida, Gainesville, Florida, United States of America; Paul E Bankey, University of Rochester, Rochester, New York, United States of America; Timothy R Billiar, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, United States of America; Steven E Calvano, University of Medicine and Dentistry of New Jersey, New Brunswick, New Jersey, United States of America; David G Camp II, Pacific Northwest National Laboratory, Richland, Washington, United States of America; Irshad H Chaudry, University of Alabama at Birmingham, Birmingham, Alabama, United States of America; Ronald W Davis, Stanford University, Palo Alto, California, United States of America; Asit K De, University of Rochester, Rochester, New York, United States of America; Bradley Freeman, Washington University, St. Louis, Missouri, United States of America; Richard L Gamelli, Loyola University Stritch School of Medicine, Maywood, Illinois, USA; Nicole S Gibran, University of Washington, Seattle, Washington, United States of America; Jeffrey L Johnson, University of Colorado Medical School, Denver, Colorado, United States of America; Matthew B Klein, Loyola University Stritch School of Medicine, Maywood, Illinois, United States of America; James A Lederer, Brigham & Women's Hospital, Boston, Massachusetts, United States of America; Stephen F Lowry, University of Medicine and Dentistry of New Jersey, New Brunswick, New Jersey, United States of America; John A Mannick, Loyola University Stritch School of Medicine, Maywood, Illinois, United States of America; Grace P McDonald-Smith, Massachusetts General Hospital, Boston, Massachusetts, United States of America; Carol L Miller-Graziano, University of Rochester, Rochester, New York, United States of America; Michael N Mindrinos, Stanford University, Palo Alto, California, United States of America; Joseph P Minei, University of Texas Southwestern Medical School, Dallas, Texas, United

States of America; Avery B Nathens, St. Michael's Hospital, Toronto, Ontario, Canada; Grant E O'Keefe, University of Washington, Seattle, Washington, United States of America; Laurence G Rahme, Massachusetts General Hospital, Boston, Massachusetts, United States of America; Daniel G Remick, Jr., Boston University School of Medicine, Boston, Massachusetts, United States of America; Michael B Shapiro, Northwestern University, Chicago, Illinois, USA; Geoffrey M Silver, Loyola University Stritch School of Medicine, Maywood, Illinois, United States of America; Richard D Smith, Pacific Northwest National Laboratory, Richland, Washington, United States of America; John D Storey, University of Washington, Seattle, Washington, United States of America; Mehmet Toner, Massachusetts General Hospital, Boston, Massachusetts, United States of America; Michael A West, Northwestern University, Chicago, Illinois, USA; Wenzhong Xiao, Stanford University, Palo Alto, California, United States of America.

This work was entirely supported by National Institutes of Health (NIH NIGMS Glue Grant 2 U54 GM062119). The principal investigator on this grant is RG Tompkins, Department of Surgery, Massachusetts General Hospital, Boston, Massachusetts, United States of America.

## DISCLOSURES

In accordance with institutional policies, authors HS Warren, C Elson, DL Hayden, DA Schoenfeld, and RG Tompkins have applied for patent protection for the method of computing summary genomic scores presented in this paper (United States Patent Application 60/858,617). The remaining authors have no financial interest in this work.

## REFERENCES

- [No author listed]. (2006) Spending on trauma care doubles to \$71.6 billion, surpassing heart disease and cancer expenditures. *Hosp. Health Netw.* 80:67.
- Meel BL. (2004) Incidence and patterns of violent and/or traumatic deaths between 1993 and 1999 in the Transkei region of South Africa. *J. Trauma* 57:125–9.

- Meislin H, *et al.* (1997) Fatal trauma: the modal distribution of time to death is a function of patient demographics and regional resources. *J. Trauma* 43:433–40.
- Suter P, *et al.* (1994) Predicting outcome in ICU patients. 2nd European Consensus Conference in Intensive Care Medicine. *Intensive Care Med.* 20:390–7.
- DeCamp MM, Demling RH. (1988) Posttraumatic multisystem organ failure. *JAMA* 260:530–4.
- Marshall JC, Vincent JL, Sibbald WJ. (1995) Descriptors of organ system dysfunction for the multiple organ dysfunction syndrome (MODS). In: *Clinical Trials for the Treatment of Sepsis*. Vincent JL, Sibbald WJ (eds.) Springer-Verlag, Berlin, pp. 122–38.
- Giannoudis PV. (2003) Current concepts of the inflammatory response after major trauma: an update. *Injury* 34:397–404.
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE. (1985) APACHE II: a severity of disease classification system. *Crit. Care Med.* 13:818–29.
- Knaus WA, *et al.* (1991) The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 100:1619–36.
- [No author]. (1998) *The Abbreviated Injury Scale 1990 Revision UPDATE* 98. Des Plaines, IL: Association for the Advancement of Automotive Medicine. 68 pp.
- Baker SP, O'Neill B, Haddon W, Jr., Long WB. (1974) The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J. Trauma* 14:187–96.
- Osler T, Baker SP, Long W. (1997) A modification of the injury severity score that both improves accuracy and simplifies scoring. *J. Trauma* 43:922–5.
- Marshall JC, Reinhart K, Eylich K, Sprung CL. (1994) A scoring system for the multiple organ dysfunction syndrome. In: *Sepsis: Current perspectives in Pathophysiology and Therapy*. Reinhart K, Eylich K, Sprung CL, (eds.) Springer-Verlag, Berlin, pp. 38–49.
- Marshall JC, *et al.* (1995) Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit. Care Med.* 23:1638–52.
- Sauaia A, *et al.* (1994) Early predictors of postinjury multiple organ failure. *Arch. Surg.* 129:39–45.
- Sauaia A, Moore FA, Moore EE, Lezotte DC. (1996) Early risk factors for postinjury multiple organ failure. *World J. Surg.* 20:392–400.
- Maier R, Bankey P, McKinley B, Freeman B. (2005) Inflammation and the Host Response to Injury, a Large-Scale Collaborative Project: Patient-Oriented Research Core-Standard Operating Procedures for Clinical Care: Foreword. *J. Trauma* 59:762–763.
- Cobb JP, *et al.* (2005) Application of genome-wide expression analysis to human health and disease. *Proc. Natl. Acad. Sci. U. S. A.* 102:4801–6.
- Feezor RJ, *et al.* (2004) Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol. Genomics* 19:247–54.

20. Common references for publications [Internet]. c2004-2007. Boston: Massachusetts General Hospital; [cited 2009 May 15]. Available from: <http://www.gluegrant.org/commonlyreferencedpubs.htm>.
21. Li C, Wong WH. (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. U. S. A.* 98:31–6.
22. Polderman KH, Jorna EM, Girbes AR. (2001) Inter-observer variability in APACHE II scoring: effect of strict guidelines and training 2. *Intensive Care Med.* 27:1365–9.
23. Polderman KH, Christiaans HM, Wester JP, Spijkstra JJ, Girbes AR. (2001) Intra-observer variability in APACHE II scoring 1. *Intensive Care Med.* 27:1550–2.
24. Chen L, *et al.* (2005) Hepatic gene expression discriminates responders and nonresponders in treatment of chronic hepatitis C viral infection. *Gastroenterology* 128:1437–44.
25. Huang E, *et al.* (2003) Gene expression predictors of breast cancer outcomes. *Lancet* 361:1590–6.